



OIKE

RAW SEQUENCE LISTING

DATE: 01/18/2002

PATENT APPLICATION: US/10/006,190

TIME: 09:23:49

Input Set : N:\Crf3\RULE60\10006190.txt

Output Set: N:\CRF3\01182002\J006190.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Hillman, Jennifer L.

6 Shah, Purvi

8 (ii) TITLE OF INVENTION: NOVEL MITOCHONDRIAL ADENYLATE KINASE

10 (iii) NUMBER OF SEQUENCES: 5

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

14 (B) STREET: 3174 Porter Drive

15 (C) CITY: Palo Alto

16 (D) STATE: CA

17 (E) COUNTRY: USA

18 (F) ZIP: 94304

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Diskette

22 (B) COMPUTER: IBM Compatible

23 (C) OPERATING SYSTEM: DOS

24 (D) SOFTWARE: FastSEQ for Windows Version 2.0

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/10/006,190

C--> 28 (B) FILING DATE: 04-Dec-2001

29 (C) CLASSIFICATION:

31 (vii) PRIOR APPLICATION DATA:

32 (A) APPLICATION NUMBER: 08/829,027

33 (B) FILING DATE:

35 (viii) ATTORNEY/AGENT INFORMATION:

36 (A) NAME: Billings, Lucy J.

37 (B) REGISTRATION NUMBER: 36,749

38 (C) REFERENCE/DOCKET NUMBER: PF-0256 US

40 (ix) TELECOMMUNICATION INFORMATION:

41 (A) TELEPHONE: 415-855-0555

42 (B) TELEFAX: 415-845-4166

43 (C) TELEX:

46 (2) INFORMATION FOR SEQ ID NO: 1:

48 (i) SEQUENCE CHARACTERISTICS:

49 (A) LENGTH: 227 amino acids

50 (B) TYPE: amino acid

51 (C) STRANDEDNESS: single

52 (D) TOPOLOGY: linear

54 (vii) IMMEDIATE SOURCE:

55 (A) LIBRARY: Consensus

56 (B) CLONE: 2122022

58 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ENTERED

RAW SEQUENCE LISTING

DATE: 01/18/2002

PATENT APPLICATION: US/10/006,190

TIME: 09:23:49

Input Set : N:\Crf3\RULE60\10006190.txt

Output Set: N:\CRF3\01182002\J006190.raw

```

60 Met Gly Ala Ser Ala Arg Leu Leu Arg Ala Val Ile Met Gly Ala Pro
61 1 5 10 15
62 Gly Ser Gly Lys Gly Thr Val Ser Ser Arg Ile Thr Thr His Phe Glu
63 20 25 30
64 Leu Lys His Leu Ser Ser Gly Asp Leu Leu Arg Asp Asn Met Leu Arg
65 35 40 45
66 Gly Thr Glu Ile Gly Val Leu Ala Lys Ala Phe Ile Asp Gln Gly Lys
67 50 55 60
68 Leu Ile Pro Asp Asp Val Met Thr Arg Leu Ala Leu His Glu Leu Lys
69 65 70 75 80
70 Asn Leu Thr Gln Tyr Ser Trp Leu Leu Asp Gly Phe Pro Arg Thr Leu
71 85 90 95
72 Pro Gln Ala Glu Ala Leu Asp Arg Ala Tyr Gln Ile Asp Thr Val Ile
73 100 105 110
74 Asn Leu Asn Val Pro Phe Glu Val Ile Lys Gln Arg Leu Thr Ala Arg
75 115 120 125
76 Trp Ile His Pro Ala Ser Gly Arg Val Tyr Asn Ile Glu Phe Asn Pro
77 130 135 140
78 Pro Lys Thr Val Gly Ile Asp Asp Leu Thr Gly Glu Pro Leu Ile Gln
79 145 150 155 160
80 Arg Glu Asp Asp Lys Pro Glu Thr Val Ile Lys Arg Leu Lys Ala Tyr
81 165 170 175
W--> 82 Glu Asp Gln Thr Lys Xaa Val Leu Xaa Tyr Tyr Gln Lys Lys Gly Val
83 180 185 190
84 Leu Glu Thr Phe Ser Gly Thr Glu Thr Asn Lys Ile Trp Pro Tyr Val
85 195 200 205
86 Tyr Ala Phe Leu Gln Thr Lys Val Pro Gln Arg Ser Gln Lys Ala Ser
87 210 215 220
88 Val Thr Pro
89 225

```

91 (2) INFORMATION FOR SEQ ID NO: 2:

93 (i) SEQUENCE CHARACTERISTICS:

94 (A) LENGTH: 854 base pairs

95 (B) TYPE: nucleic acid

96 (C) STRANDEDNESS: single

97 (D) TOPOLOGY: linear

99 (vii) IMMEDIATE SOURCE:

100 (A) LIBRARY: Consensus

101 (B) CLONE: 2122022

103 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

105 GCCANGCCCA AAGCCCTGGT ACCCGCGCGG TGGGGCCTCA GTCTGCGGCC ATGGGGGCGT 60
106 CCGCGCGGCT GCTGCGAGCG GTGATCATGG GGGCCCCGGG CTCGGGCAAG GGCACCGTGT 120
107 CGTCGCGCAT CACTACACAC TTCGAGCTGA AGCACCTCTC CAGCGGGGAC CTGCTCCGGG 180
108 ACAACATGCT GCGGGGCACA GAAATTGGCG TGTAGCCAA GGCTTTCATT GACCAAGGGA 240
109 AACTCATCCC AGATGATGTC ATGACTCGGC TGGCCCTTCA TGAGCTGAAA AATCTCACCC 300
110 AGTATAGCTG GCTGTTGGAT GGTTTTCCAA GGACACTTCC ACAGGCAGAA GCCCTAGATA 360
111 GAGCTTATCA GATCGACACA GTGATTAACC TGAATGTGCC CTTTGAGGTC ATTAAACAAC 420
112 GCCTTACTGC TCGCTGGATT CATCCGCCA GTGGCCGAGT CTATAACATT GAATTCAACC 480
113 CTCCCAAAC TGTGGGCATT GATGACCTGA CTGGGGAGCC TCTCATTCAG CGTGAGGATG 540

```

RAW SEQUENCE LISTING

DATE: 01/18/2002

PATENT APPLICATION: US/10/006,190

TIME: 09:23:49

Input Set : N:\Crf3\RULE60\10006190.txt

Output Set: N:\CRF3\01182002\J006190.raw

```

114 ATAAACCAGA GACGGTTATC AAGAGACTAA AGGCTTATGA AGACCAAACA AAGNCAGTCC 600
115 TGGNATATTA CCAGAAAAAA GGGGTGCTGG AAACATTCTC CGGAACAGAA ACCAACAAGA 660
116 TTTGGCCCTA TGTATATGCT TTCCTACAAA CTAAAGTTCC ACAAAGAAGC CAGAAAGCTT 720
117 CAGTTACTCC ATGAGGAGAA ATGTGTGTAA CTATTAATAG TAAGATGGGC AAACCTCCTA 780
118 GTCCTTGCAT TTAGAAGCTG CTTTTCCTAA GACTTCTAGT ATGTATGAAT TCTTTGAAAA 840
119 TTATATTACT TTTA 854

```

121 (2) INFORMATION FOR SEQ ID NO: 3:

123 (i) SEQUENCE CHARACTERISTICS:

```

124 (A) LENGTH: 227 amino acids
125 (B) TYPE: amino acid
126 (C) STRANDEDNESS: single
127 (D) TOPOLOGY: linear

```

129 (vii) IMMEDIATE SOURCE:

```

130 (A) LIBRARY: GenBank
131 (B) CLONE: 217576

```

133 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

135 Met Gly Ala Ser Ala Arg Leu Leu Arg Ala Ala Ile Met Gly Ala Pro
136 1 5 10 15
137 Gly Ser Gly Lys Gly Thr Val Ser Ser Arg Ile Thr Lys His Phe Glu
138 20 25 30
139 Leu Lys His Leu Ser Ser Gly Asp Leu Leu Arg Asp Asn Met Leu Arg
140 35 40 45
141 Gly Thr Glu Ile Gly Val Leu Ala Lys Thr Phe Ile Asp Gln Gly Lys
142 50 55 60
143 Leu Ile Pro Asp Asp Val Met Thr Arg Leu Val Leu His Glu Leu Lys
144 65 70 75 80
145 Asn Leu Thr Gln Tyr Asn Trp Leu Leu Asp Gly Phe Pro Arg Thr Leu
146 85 90 95
147 Pro Gln Ala Glu Ala Leu Asp Arg Ala Tyr Gln Ile Asp Thr Val Ile
148 100 105 110
149 Asn Leu Asn Val Pro Phe Glu Val Ile Lys Gln Arg Leu Thr Ala Arg
150 115 120 125
151 Trp Ile His Pro Gly Ser Gly Arg Val Tyr Asn Ile Glu Phe Asn Pro
152 130 135 140
153 Pro Lys Thr Met Gly Ile Asp Asp Leu Thr Gly Glu Pro Leu Val Gln
154 145 150 155 160
155 Arg Glu Asp Asp Arg Pro Glu Thr Val Val Lys Arg Leu Lys Ala Tyr
156 165 170 175
157 Glu Ala Gln Thr Glu Pro Val Leu Glu Tyr Tyr Arg Lys Lys Gly Val
158 180 185 190
159 Leu Glu Thr Phe Ser Gly Thr Glu Thr Asn Lys Ile Trp Pro His Val
160 195 200 205
161 Tyr Ala Phe Leu Gln Thr Lys Leu Pro Gln Arg Ser Gln Glu Thr Ser
162 210 215 220
163 Val Thr Pro
164 225

```

166 (2) INFORMATION FOR SEQ ID NO: 4:

168 (i) SEQUENCE CHARACTERISTICS:

```

169 (A) LENGTH: 227 amino acids

```

RAW SEQUENCE LISTING

DATE: 01/18/2002

PATENT APPLICATION: US/10/006,190

TIME: 09:23:49

Input Set : N:\Crif3\RULE60\10006190.txt

Output Set: N:\CRF3\01182002\J006190.raw

```

170      (B) TYPE: amino acid
171      (C) STRANDEDNESS: single
172      (D) TOPOLOGY: linear
174      (vii) IMMEDIATE SOURCE:
175          (A) LIBRARY: GenBank
176          (B) CLONE: 450312
178      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
180 Met Gly Ala Ser Gly Arg Leu Leu Arg Ala Val Ile Met Gly Ala Pro
181 1      5      10      15
182 Gly Ser Gly Lys Gly Thr Gly Ser Ser Arg Ile Thr Lys His Phe Glu
183      20      25      30
184 Leu Lys His Leu Ser Ser Gly Asp Leu Leu Arg Gln Asn Met Leu Gln
185      35      40      45
186 Gly Thr Glu Ile Ala Val Leu Ala Lys Ser Phe Ile Asp Gln Gly Lys
187      50      55      60
188 Leu Ile Pro Asp Asp Asp Met Thr Arg Leu Ala Leu His Glu Leu Lys
189      65      70      75      80
190 Asn Leu Thr Gln Cys Ser Trp Leu Leu Asp Gly Phe Pro Arg Thr Leu
191      85      90      95
192 Pro Gln Ala Glu Ala Leu Asp Arg Val Tyr Gln Ile Asp Thr Val Ile
193      100     105     110
194 Asn Leu Asn Val Pro Phe Glu Val Ile Lys Leu Arg Leu Thr Ala Arg
195      115     120     125
196 Trp Ile His Pro Ala Ser Gly Arg Val Tyr Asn Ile Glu Phe Asn Pro
197      130     135     140
198 Pro Lys Thr Val Gly Ile Asp Asp Leu Thr Gly Glu Pro Leu Ile Gln
199      145     150     155     160
200 Arg Glu Asp Asp Lys Pro Glu Thr Val Ile Lys Arg Leu Lys Ala Tyr
201      165     170     175
202 Glu Ala Gln Thr Glu Pro Val Leu Gln Tyr Tyr Gln Lys Lys Gly Val
203      180     185     190
204 Leu Glu Thr Phe Ser Gly Thr Glu Thr Asn Lys Ile Arg Pro His Val
205      195     200     205
206 Tyr Ser Phe Leu Gln Met Lys Val Pro Glu Thr Ile Gln Lys Ala Ser
207      210     215     220
208 Val Thr Pro
209 225
211 (2) INFORMATION FOR SEQ ID NO: 5:
213      (i) SEQUENCE CHARACTERISTICS:
214          (A) LENGTH: 223 amino acids
215          (B) TYPE: amino acid
216          (C) STRANDEDNESS: single
217          (D) TOPOLOGY: linear
219      (vii) IMMEDIATE SOURCE:
220          (A) LIBRARY: GenBank
221          (B) CLONE: 28577
223      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
225 Met Ala Ser Lys Leu Leu Arg Ala Val Ile Leu Gly Pro Pro Gly Ser
226 1      5      10      15

```

RAW SEQUENCE LISTING

DATE: 01/18/2002

PATENT APPLICATION: US/10/006,190

TIME: 09:23:49

Input Set : N:\Crf3\RULE60\10006190.txt

Output Set: N:\CRF3\01182002\J006190.raw

```

227 Gly Lys Gly Thr Val Cys Gln Arg Ile Ala Gln Asn Phe Gly Leu Gln
228           20           25           30
229 His Leu Ser Ser Gly His Phe Leu Arg Glu Asn Ile Lys Ala Ser Thr
230           35           40           45
231 Glu Val Gly Glu Met Ala Lys Gln Tyr Ile Glu Lys Ser Leu Leu Val
232           50           55           60
233 Pro Asp His Val Ile Thr Arg Leu Met Met Ser Glu Leu Glu Asn Arg
234 65           70           75           80
235 Arg Gly Gln His Trp Leu Leu Asp Gly Phe Pro Arg Thr Leu Gly Gln
236           85           90           95
237 Ala Glu Ala Leu Asp Lys Ile Cys Glu Val Asp Leu Val Ile Ser Leu
238           100          105          110
239 Asn Ile Pro Phe Glu Thr Leu Lys Asp Arg Leu Ser Arg Arg Trp Ile
240           115          120          125
241 His Pro Pro Ser Gly Arg Val Tyr Asn Leu Asp Phe Asn Pro Pro His
242           130          135          140
243 Val His Gly Ile Asp Asp Val Thr Gly Glu Pro Leu Val Gln Gln Glu
244 145           150           155           160
245 Asp Asp Lys Pro Glu Ala Val Ala Ala Arg Leu Arg Gln Tyr Lys Asp
246           165           170           175
247 Val Ala Lys Pro Val Ile Glu Leu Tyr Lys Ser Arg Gly Val Leu His
248           180          185          190
249 Gln Phe Ser Gly Thr Glu Thr Asn Lys Ile Trp Pro Tyr Val Tyr Thr
250           195          200          205
251 Leu Phe Ser Asn Lys Ile Thr Pro Ile Gln Ser Lys Glu Ala Tyr
252           210          215          220

```

VERIFICATION SUMMARY

DATE: 01/18/2002

PATENT APPLICATION: US/10/006,190

TIME: 09:23:50

Input Set : N:\Crf3\RULE60\10006190.txt

Output Set: N:\CRF3\01182002\J006190.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:82 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1